AD_____

Award Number: DAMD17-00-1-0365

TITLE: A Novel Strategy to Isolate Invasion-Inducing Proteins From Human Breast Tumors

PRINCIPAL INVESTIGATOR: Ian P. Whitehead, Ph.D.

CONTRACTING ORGANIZATION: University of Medicine and Dentistry
New Jersey Medical School
Newark, New Jersey 07103-2714

REPORT DATE: June 2001

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2001	Annual (1 Jun 00 - 31 May 01)

4. TITLE AND SUBTITLE
A Novel Strategy to Isolate Invasion-Inducing Proteins From Human Breast Tumors

5. FUNDING NUMBERS
DAMD17-00-1-0365

6. AUTHOR(S)
Ian P. Whitehead, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Medicine and Dentistry, New Jersey Medical School
Newark, New Jersey 07103-2714

E-Mail: whiteip@UMDNJ.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

This proposal describes the development of a screening strategy that can be applied to the identification of genes that regulate the progression of breast tumors to a metastatic state. The general approach is to make expressible cDNA libraries from highly invasive breast tumor cell lines, and to transfer these libraries (and the phenotype) to human tumor cells that normally do not exhibit invasive properties. This report covers the first 10 months of this project and describes (i) the construction and characterization of 6 retrovirus-based expression libraries, (ii) the efficient conversion of plasmid-based libraries into libraries of high titre infectious retroviral particles, and (iii) a preliminary test screen of one of the libraries. A population of 250,000 infected cells has been passed through three rounds of enrichment and analyzed for the presence of proviral inserts. A large, but manageable number of inserts have been identified by PCR and are now being recovered for further analysis. To summarize, the screening is progressing according to our original schedule, and we anticipate that some of the proviral inserts that we are currently recovering will exhibit invasive potential.

14. SUBJECT TERMS
Invasion, Metastasis, Expression Cloning, Signaling, Oncogenes

15. NUMBER OF PAGES
25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover.....1

SF 298.....2

Table of Contents.....3

Introduction.....4

Body.....4-7

Key Research Accomplishments.....7

Reportable Outcomes.....8

Conclusions.....8

References.....8

Appendices.....9-17

INTRODUCTION

Retrovirus-based cDNA expression libraries provide a powerful tool for the identification of oncogenes (1-3). In this proposal, an expression system that has been developed by the principle investigator (2) has been adapted for the isolation of genes that contribute to the progression of breast carcinoma to a metastatic state. Since the success of this proposal critically depends upon our ability to generate large populations of target cells that stably express complex libraries, a substantial amount of time in the first ten months was allocated for protocol development and evaluation. This allocation of time and effort is consistent with what was proposed in the "statement of work" that was submitted with the initial proposal. The refined set of protocols that was developed for use in this study is described in detail in an recent volume of "Methods in Enzymology (4; see appendix 1). The work that was conducted during the project period for this report can be broken down into three phases: (i) construction and evaluation of complex cDNA expression libraries from breast tumor-derived cell lines, (ii) isolation of mass populations of retroviral-infected target cells and (iii) preliminary screening of an MBA-MB-231 derived library. Each of these steps is described in detail below.

BODY

Construction and evaluation of cDNA expression libraries. The vast majority of transforming cDNAs that have been identified to date have been isolated in a truncated or mutated form. This suggests that most oncogenes arise as a consequence of mutational activation, rather than through deregulated expression of the wild-type sequences. Thus, one can reasonably presume that the representation of an oncogenic cDNA within a library population may be somewhat lower than the representation of the wild-type cDNA. From the perspective of cDNA library construction this means that the successful identification of oncogenes may increase dramatically as the complexity of the library increases. A more complex library is more likely to express a low abundance, mutated cDNA than a less complex library. With this in mind, a considerable amount of time and effort has been devoted to the production of complex libraries derived from highly invasive, breast tumor-derived cell lines. The factors that have the most dramatic impact upon successful library construction include the quality of RNA, the efficiency of cDNA synthesis, efficient incorporation of cDNAs into retroviral vectors, and efficient library amplification (2). All of these parameters were re-evaluated and a new generation of library construction protocols was developed. These protocols are described in detail in a recent edition of "Methods in Enzymology" (2; see appendix 1).

Once the library construction protocols were optimized, a series of libraries was constructed using RNAs isolated from the MBA-MB-231 and MBA-MB-435 cell lines (5-6) as described in Specific Aim I of the proposal. cDNAs were size fractionated into three populations, and three libraries were constructed for each cell line based on these fractions. The total number of cDNAs in each library was determined using methods that have been described previously (2), as was the percentage of clones that contained inserts, the average insert size, and the size range of inserts. This data is summarized

in Table 1. Based on our previous experience with library construction (2), we are pleased with the complexities of the libraries which range from a high of 950,000 clones down to a low of 265,000 clones. Consistent with what we have previously observed, libraries derived from the smaller RNA fractions tend to have a higher degree of complexity because of the greater ease in cloning smaller cDNAs. In the case of all the libraries, greater than 90% of the clones that were examined contained cDNA inserts and the average size of these inserts corresponded well with the size fraction from which the library was derived. To summarize, we consider libraries described in Table 1 to be of high quality and suitable for the screens that we are conducting.

Table 1: Retroviral-based cDNA expression libraries derived from human breast tumor cell lines.

Designation	# of Clones[a]	% Inserts[b]	Range (kb)	Average Length (kb)
MBA-MB-231-S	950,000	97	1.0 – 2.5	1.55
MBA-MB-231-M	575,000	90	2.5 – 3.5	2.70
MBA-MB-231-L	325,000	98	3.5 – 6.0	4.15
MBA-MB-435-S	810,000	97	1.1 – 2.9	1.35
MBA-MB-435-M	615,000	93	1.9 – 3.1	2.65
MBA-MB-435-L	265,000	91	3.1 – 5.7	4.30

[a] Estimated number of clones contained within each library. Estimate is based on manual counts of colonies on representative plates following library electroporation.

[b] Following electroporation, 50-100 library colonies are chosen at random from a plate. Plasmids are recovered from each colony and the % of plasmids containing inserts, the range of insert size and the average insert size are determined.

Establishment of mass populations of MCF-7 cells that have been stably infected with the retroviral-based libraries. The second component of Specific Aim 1 requires that we convert our libraries of retroviral particles into large, representative populations of stably infected cells. Optimally we wished to achieve a conversion ratio (no. of infecting cDNAs / no. of recipient cells) of 1 in order to maintain the complexity of the libraries. In our proposal, we described the construction of an MCF-7 human breast epithelial cell line cell line that stably expresses the ecotropic receptor. This cell line gives us the ability to package our libraries with the more efficient ecotropic packaging cell lines (such as BOSC23), yet conduct our screens in a human, breast-derived cell line. Using cDNAs derived from our expression libraries we wished to test these cells to determine the viral titres that we could achieve, and then compare this to the highly infectable NIH 3T3 cells. BOSC23 cells were transiently transfected with library DNA and then viral titres were determined by infection of the modified MCF-7 cells or NIH 3T3 cells, followed by selection in hygromycin. Using this system we found that that we are consistently able to achieve viral titres in excess of 1×10^7 viral particles / ml of media (Table 2). This frequency of infection was only marginally lower that what we

were able to achieve using the highly infectable murine NIH 3T3 cell line as a recipient. Since NIH 3T3 cells can be infected at frequencies greater than 95% (2), we presume that we are achieving similar infection frequencies with our modified MCF-7 cell line. The library titres that we have achieved are consistently higher than what we have achieved in past screens and should be sufficiently high to permit very comprehensive screens of our libraries.

Table 2: Stable expression of the ecotropic receptor renders MCF-7 cells susceptible to high frequency of infection by libraries packaged in an ecotropic cell line (BOSC23)[a]

Packaging Cells	Recipient Cells	Library	Infection Rate[b]
BOSC23	NIH 3T3	MDA-MB-231-S	1.0×10^7 1.6×10^7 1.3×10^7
		MBA-MB-435-S	2.1×10^7 1.9×10^7 1.2×10^7
BOSC23	MCF-7-Eco^R	MBA-MB-231-S	1.5×10^7 1.3×10^7 9.5×10^6
		MBA-231-435-S	1.1×10^7 1.3×10^7 1.3×10^7

[a] Comparison of retrovirus infection efficiencies achieved with different recipient cell lines. Six-centimeter culture dishes containing 1.5×10^6 packaging cells were transfected with 10 ug of library DNA utilizing standard calcium transfection procedures. Secreted viral soup was collected at 48 hr post-transfection and assayed on 10 cm dishes containing 5×10^5 recipient cells. NIH 3T3 is a murine fibroblast cell line that is highly infectable by ecotropic virus and which was included as a positive control for infection. MCF-7-Eco^R is a cell line that has been stably transfected with the ecotropic receptor. Each data point shown is from a single experiment performed on triplicate plates.

[b] Number of hygromycin resistant colonies that were generated per ml of viral soup. Values are provided for three independent experiments.

Test screen of the MBA-MB-231-S library: Having constructed a series of complex libraries, and having established protocols to screen such libraries, a test screen was initiated using the MBA-MB-231-S library. The purpose of this screen was to determine whether serial passaging of infected populations of cells through transwell based invasion filters would generate populations of cells that were enriched for a finite, manageable number of proviral inserts as determined by PCR amplification. BOSC23 cells were transfected with 10 ug of DNA from the MBA-MB-231-S library. At 48 hr, 10 ml of viral soup was collected and used to infect 1×10^6 MCF-7-Eco^R cells. Cells were

placed under hygromycin selection and allowed to grow to confluence. Cells were then trypsinized, counted (approximately 250,000 cells), and then divided into 5 subpopulations, each of which was passed through a collagen stacked transwell filter. After 24 hr, cells that had passed through the filter were harvested, replated and grown to confluence. Due to the low number of cells that passed through the filter (an indication of the low background associated with this screen), amplification of the invasive cell population took approximately 3-4 weeks. Cells were then trypsinized and passed through a new filter, and the enrichment was repeated. Following the third enrichment (at week 12), genomic DNA was isolated from the cell clone and a PCR reaction was performed using primers that would specifically recognize proviral inserts. Discrete bands were amplified from each of the 5 samples as shown in Figure 1. Interestingly, a similar size band of approximately 2.0 kB was independently amplified in four of the five samples. These bands will now be recovered using methods that have been described previously. The ability to enrich our cell population to the point were only 2 – 5 proviral inserts remain after three cycles of selection suggests that larger scale screens can be performed in a relatively time-efficient manner.

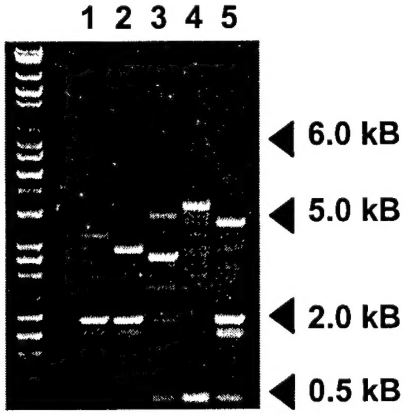

Figure 1: PCR amplification of genomic DNA isolated from mass populations of cells that have been serially enriched for invasive clones. MCF-7-EcoR cells were stably infected with retroviral particles derived from the MBA-MB-231-S library. Five cell populations (50,000 cells each) were screened for invasive clones. PCR bands indicate the presence of proviral inserts within the cell populations. DNA size markers are indicated on the right.

KEY RESEARCH ACCOMPLISHMENTS

- Construction of highly complex cDNA expression libraries from RNA derived from invasive human breast tumor cell lines.
- Establishment of mass populations of human cells that are stably infected with retroviral-based cDNA expression libraries
- Establishment of enrichment protocols that will be required to screen libraries for cDNAs whose expression contribute to an invasive phenotype.

REPORTABLE OUTCOMES

Manuscripts:

G. M. Mahon and I. P. Whitehead (2001) Retrovirus cDNA expression library screening for oncogenes. Methods enzymol. 332:211-221.

CONCLUSIONS

The work performed during the first 10 months of this award has focused on several issues that must be addressed prior to the initiation of large-scale screens. For example, considerable time and attention was dedicated to the production of cDNA expression libraries that were of sufficient complexity for the proposed screens. In addition, it was necessary to determine the number of rounds of enrichment that must be performed prior to initiating the process of cDNA recovery. Because of the relatively large lag time (greater than 12 weeks) required to amplify the enriched cell populations, screens will need to be staggered to ensure an efficient use of time and manpower. This work that was described for the first 10 months in the "Statement of Work" has now been successfully completed and large-scale screening efforts have begun. Interestingly, the small pilot screens that have been conducted have already identified a proviral insert that was independently identified in several screens of the same library (four out of five). We have recently recovered this cDNA and are now testing it for its invasive potential.

REFERENCES

1. T. Kitamura, M. Onishi, S. Kinoshita, A. Shibuya, A. Miyajima, and G. P. Nolan (1995) Efficient screening of retroviral cDNA expression libraries. *Proc. Natl. Acad. Sci. (U S A)* **92,** 9146-50.

2. I. Whitehead, H. Kirk, and R. Kay (1995) Expression cloning of oncogenes by retroviral transfer of cDNA libraries. *Mol. Cell. Biol.* **15,** 704-10.

3. G. J. Hannon, P. Sun, A. Carnero, L. Y. Xie, R. Maestro, D. S. Conklin, and D. Beach (1999) MaRX: an approach to genetics in mammalian cells. *Science* **283,** 1129-30.

4. G. M. Mahon and I. P. Whitehead (2001) Retrovirus cDNA expression library screening for oncogenes. Methods enzymol. 332:211-221.

5. Cailleau R, Young R, Olive M, and Reeves WJ Jr. 1974 Breast tumor cell lines from pleural effusions. *J. Natl. Cancer Inst.* **53**, 661-674.

6. Cailleau R, Olive M, and Cruciger QV. 1978 Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In vitro* **14**, 911-915.

APPENDIX 1

Reprints of Manuscripts

G. M. Mahon and I. P. Whitehead (2001) Retrovirus cDNA expression library screening for oncogenes. Methods enzymol. 332:211-221.

Methods in Enzymology 2001; 332: 211-21

Retrovirus cDNA Expression Library Screening for Novel Oncogenes

Running head: Retroviral Screens for Oncogenes

Gwendolyn M. Mahon and Ian P. Whitehead*

Department of Microbiology and Molecular Genetics

MSB F607, UMDNJ-New Jersey Medical School

185 South Orange Ave, Newark, NJ 07103-2714

* Send proofs to

Although conventional, plasmid-based, gene transfer assays have been used to identify proto-oncogenes, only a handful of transforming sequences have been detected in this manner. The limited success of these efforts reflects technical limitations to the systems employed, rather than exhaustion of the pool of cDNAs with oncogenic potential. The most severe limitations associated with the use of plasmid-based libraries for expression cloning (low efficiencies of transfection and expression) have now been overcome by the development of retroviral-based cDNA transfer systems.[1-4] In these systems, cDNA expression libraries are constructed in retroviral plasmids, and then converted into libraries of infectious retroviral particles. Four major advantages are obtained through the use of retroviral library transfer: (i) the ability to screen large numbers of cDNA clones on an equivalent number of recipient cells, (ii) the relatively high levels of expression obtained with retrovirally transferred cDNAs, (iii) the potential to use cell lines which have been inaccessible to expression cloning because of low transfection efficiencies, and (iv) the development of highly efficient recovery mechanisms for the proviral inserts.

We have described a retroviral-based cDNA expression system and its successful application to the identification of novel oncogenes.[2] This system, which is described here in detail, permits the stable transfer and expression of large numbers of cDNA clones into equivalent numbers of recipient cells. This allows for the efficient screening of complex cDNA libraries, and facilitates the identification of transforming sequences that are present at very low frequency within the cDNA population. Although we are describing methods that identify transforming cDNAs based on their focus-forming activity,[5] this expression system can be readily adapted for use in alternative screens for growth transformation.

Construction of retroviral-based cDNA expression libraries:

Retroviral vectors: We have described the construction of a retroviral vector (pCTV3B) that has been designed specifically for use in cDNA expression cloning (Fig. 1).[2] The vector has been made as compact as possible (≈5560 bp) to maximize the stability of cDNAs, and all splice sites have been removed so that they cannot activate cryptic splice sites that may be present within the cDNA inserts. In addition to viral sequences, the viral transcripts that are derived from pCTV3B contain an inserted cDNA, a bacterial selectable marker (*supF*),[6] and a marker that permits selection in mammalian cells (Hygro^r).[7] An extended *gag* region ensures efficient packaging of this viral transcript when it is expressed in an appropriate packaging cell line.[8]

The pCTV retroviral vectors utilize the *supF*-p3 selection system for *Escherichia coli* (*E. coli*) that was developed by Brian Seed.[9] The *supF* gene encodes a tRNA molecule that can rescue amber mutations within antibiotic resistance genes, and thus provides a compact (≈220 bp) bacterial selectable marker for the retroviral vectors. The p3 plasmid is a 50 kbp, single copy, stably propagated plasmid that encodes a wild-type kanamycin (Kan^r) resistance gene as well as amber mutant ampicillin and tetracycline resistance genes. *E. coli* strains that contain p3 (e.g. MC1061/p3[10]) can be used to select for plasmids that are carrying *supF*.

Isolation of poly(A) mRNA: The successful construction of a complex library (>10^6 clones) generally requires 3-5 μg of good quality poly(A) mRNA. Such quantities can be readily obtained from tumor-derived, mammalian cell lines. Total RNA is first isolated using Trizol™ reagent (GIBCO/BRL) according to the manufacturers

instructions. We collect lysates from at least 2 X 10^8 cells (e.g. 20 confluent 100 mm dishes) which will yield in excess of 1 mg of total RNA. The RNA is then subjected to two rounds of purification on oligo(dT) cellulose columns to obtain 10-50 μg of purified poly(A) mRNA.[11] Since we generally use random priming during cDNA synthesis, it is important to carry out the second round of purification to remove as much ribosomal RNA as possible.

Preparation of cDNAs: When constructing cDNA libraries that are to be screened for oncogenes, we find it advantageous to use random priming and bi-directional cloning. The majority of oncogenes that have been described, to date, are activated by truncation, and random priming enriches the library for truncated cDNAs. Bi-directional cloning allows us to simultaneously screen for oncogenes and tumor suppressors. The protocol that we are describing generates blunt ended cDNAs that can be fused to *Bst*XI linkers, and then cloned into the retroviral pCTV3B vector. We use SuperScript™ II reverse transcriptase (GIBCO/BRL) for first strand synthesis. Second strand synthesis is carried out by nick translation using *E.coli* DNA polymerase I, *E. Coli* RNase H and *E.Coli* DNA ligase (all from GIBCO/BRL).

All glassware and solutions used for first strand synthesis must be RNase free. For each reaction it is preferable to work with at least 3-5 μg of good quality mRNA that has been resuspended in RNase free milliQ water. This should generate sufficient quantities of cDNA to allow size fractionation of the libraries. In a sterile 1.5 ml tube combine 4 μl of a 50 ng/μl random hexamer mix, 5 μg of mRNA, and RNase free water to a final volume of 21 μl. Incubate at 70° for 10 min, and then chill on ice. Add 8 μl of 5X First

Strand buffer (GIBCO/BRL), 4 μl of 0.1 *M* dithiothreitol (DTT), and 2 μl of a 10 m*M* dNTP mix. Mix contents gently by vortexing, and incubate for 2 min at 37°. Add 5 μl of SuperScript™ II (200 U/μl), mix gently, and continue incubation (37°) for 1 hr. Place the mixture on ice and proceed with second strand synthesis.

To the first strand reaction mixture add the following, in order (on ice): 71 μl dH_2O, 30 μl 5X Second Strand buffer (GIBCO/BRL), 3 μl 10 m*M* dNTP mix, 1μl (10 U) *E. coli* DNA ligase, 4 μl (40 U) *E. coli* DNA polymerase I, and 1 μl (2 U) *E. coli* RNase H. Mix by gently vortexing, and incubate at 16° for 2 hr. Then, add 2 μl (10 U) of T4 DNA polymerase and incubate at 16° for a further 5 min. Place on ice and stop the reaction with 10 μl of 0.5 *M* EDTA. Extract with 150 μl phenol/chloroform, and precipitate with 70 μl of 7.5 *M* NH_4OAc and 500 μl of ice cold, 100% EtOH. Centrifuge at 14,000 *g* for 20 min, wash the pellet with ice cold 70% EtOH, and resuspend in 20 μl dH_20. Load an aliquot (2 μl) on a 1.2% agarose gel to check the quality and size range of the cDNAs.

Linkering of cDNAs: The *Bst*XI linkering system that has been adapted for use with the pCTV vectors has been described previously.[9] The *Bst*XI linkers substantially decrease the number of transformants that do not contain an insert, and reduce the possibility of having two cDNAs ligated into the same vector. These adapters have also been designed to introduce stop codons in all three reading frames at the 3' end of the cDNA, thus preventing translational readthrough into the vector sequences. Oligonucleotides should be synthesized with 5' phosphates or should be kinased prior to annealing (13mer = 5'-TCA GTT ACT CAG G, 17mer = 5'-CCT GAG TAA CTG ACA CA). To anneal oligonucleotide pairs, combine equimolar amounts of each oligonucleotide, heat

to 60°, and then allow to slowly cool to room temperature. Ligation reactions consist of 1-5 μg cDNA, 0.5 μg linkers, 4 μl 5X ligation buffer, 2 U T4 DNA ligase, and water to 20 μl. Incubate the ligation reactions at 16° for 4-8 hr.

Size fractionation of cDNA: In order to avoid the selection against large cDNAs that often occurs during library construction, we have found it beneficial to separate our linkered cDNAs into several size classes prior to ligation into the retroviral vector. This facilitates the separate transformation of larger cDNAs and affords the opportunity to adjust their representation in the final libraries. Load the cDNA linkering reaction onto a 1.2 % agarose gel (containing ethidium bromide) with flanking size markers. Run the cDNA a sufficient distance into the gel to get reasonable marker separation while still minimizing the size of the gel slice. Slice off the flanking marker lanes, expose them to UV, and mark off the appropriate size ranges. Do not expose the cDNA lane to UV. We normally isolate fractions with size ranges 1000-2500 bp and 2500+ bp. Realign the marker lanes with the cDNA lane and excise the appropriate sized gel segments. Purify the cDNA using the GeneClean™ procedure (Bio 101). This provides reasonable yields of cDNA while getting rid of unligated adapters.

Ligation of linkered cDNA: The cloning site of the pCTV3B retroviral vector consists of two *Bst*XI sites separated by a 400 bp stuffer fragment (Fig. 1).[11] Complete digestion of the vector with *Bst*XI removes the stuffer and generates two non-complementary ends that can be ligated to the *Bst*XI linkered cDNAs. The retroviral vector is digested to completion with *Bst*XI, dephosphorylated, and purified on a 1.2% agarose gel followed by

electroelution and precipitation. As with the cDNA, it is important not to expose the vector DNA to UV at any time. Ligation reactions consist of: 20 ng vector, equimolar amount of cDNA, 1 μl 10X ligation buffer (250 m*M* Tris.Cl pH 7.8, 50 m*M* $MgCl_2$, 10 m*M* DTT, 10 m*M* ATP), H_2O to 10 μl total, and 0.5 U T4 Ligase. This ligation reaction includes a minimal ionic strength ligation buffer that is required specifically for electroporation. Incubate the ligation at room temperature for 3 hr, then add 10 μl H_2O and heat kill at 72° for 20 min.

Generating retroviral plasmid libraries: Libraries of retroviral plasmids are generated by electro-transformation of MC1061/p310 *E. coli* (we use a BioRad Gene Pulser II).[10] The transformation mixes are plated in soft agar which substantially reduces the variation in colony size and allows large numbers of colonies to be plated (up to 750,000) on single 15 cm plates. Since the size of the library is determined, in part, by the success of the electroporation, we only use electro-competent MC1061/p3 cells that have a competence greater than 5×10^9.

Thaw electro-competent MC1061/p3 *E. coli* on ice until completely melted. Add up to 1.5 μl of your ligation mix to a pre-chilled tube, and then add and gently mix 40 μl of electro-competent cells. Allow to sit for 1 min on ice and then transfer to a pre-chilled 1 mm cuvette. Electroporate the cells (1600V, 25μF, and 200 ohms) and then immediately add 1 ml of YT broth (with 20 m*M* Glucose) to the cuvette.[11] Transfer the cells to a new tube, incubate in a 37° water bath for 10 min, and then shake at 37° for an additional 110 min. Bring the volume of the cell solution up to 5 ml with more YT and warm to 37° in a water bath. Put 5 ml of melted 1.2% YT agar containing 100 μg/ml ampicillin and 15 μg/ml tetracycline in a second tube and incubate in a 44° water bath. Once both tubes are

temperature equilibrated, combine the two tubes, mix by inversion, and quickly pour on the surface of a 15 cm plate of 1.2% agar/YT (50 μg/ml ampicillin/7.5 μg/ml tetracycline). Let the soft agar set at room temperature for 15 min and then incubate at 37° for 20 hr. The library size can now be determined by counting the number colonies on the plates.

Recovering the plasmid library from soft agar: Once the plasmid library has been grown in soft agar, we recommend that the cells be harvested, and the DNA isolated, within 24 hr. Plates can be stored at 4° prior to harvesting. Gently wash the surface of the plate with YT to remove most of the large surface colonies. Scrape off the soft agar into 50 ml Falcon tubes, mash it well, and then transfer it to a 50 ml syringe. Add 10 ml of YT to the slurry and run it three times through an 18G needle and then three times through a 22G needle. Mix in 10 ml of Sephadex G25 (medium; SIGMA) that has been autoclaved in YT (G25/YT). Prepare spin columns (1 per plate) from 20 ml syringe barrels, plugged with a wad of glass wool and resting in 50 ml Falcon tubes. Add a 10 ml bed of G25/YT to each column and then overlay with the library/G25/YT mix. Centrifuge the columns at 500 rpm in a swinging bucket rotor for 5 min, then increase to 1000 rpm for an additional 10 min. Centrifuge the eluted medium at 4000 *g* for 10 min to recover the cells, and then resuspend the pellet in 10 volumes of YT. To prepare a glycerol stock, add 0.5 ml of the cells to 0.5 ml YT/40% glycerol, freeze in a dry ice/ethanol bath for 10 min, and then store at –70°. To prepare a library of plasmid DNA, dilute the remaining cells to A_{600} = 0.1 in YT, and then grow to A_{600} = 1.5. DNA can now be prepared by alkaline lysis,[12] followed by purification through ethidium bromide-cesium chloride gradients.[11]

Converting plasmid libraries to viral libraries: Packaging cells are mammalian cell lines that stably express all the viral components that are necessary to recognize and package viral mRNAs. Such cell lines can facilitate the simultaneous and proportional conversion of a plasmid library into a library of retroviral particles. When pooled plasmid DNAs from the cDNA library are introduced into these cell lines, viral mRNAs (including the inserted cDNA) are transcribed, packaged into infectious retroviral particles, and released into the media. These particles can then be harvested and either stored, or used for screening. Packaging cell lines vary in the viral titers that can be obtained. We have found that the high titers ($>10^6$/ml) that can be routinely obtained with the BOSC23 cell line are more than sufficient to ensure the efficient screening of highly complex libraries.[13]

Pooled plasmid DNAs from the cDNA libraries are introduced into the BOSC23 cells by calcium phosphate transfection. Unlike most adherent cell lines, BOSC23 cells do not form even monolayers and will begin clumping before confluence is reached. To ensure a high efficiency of transfection, it is important that this clumping be minimized. To achieve this, we begin splitting the cells 1:1, 2-3 days prior to transfection. If this is repeated on 2-3 consecutive days, it is possible to obtain high density plates of well spread cells. These cells can be readily trypsinized, and can be accurately counted and plated.

Plate 2 X 10^6 BOSC23 cells per plate (60 mm), 24 hr prior to transfection. Optimal transfection density is about 80% confluence. Be precise! Relatively small variations in cell density, either high or low, can sometimes have dramatic effects on transfection efficiencies and/or cell survival. Prior to transfection, change the media to 4 ml of

DMEM/10% fetal bovine serum (FBS) containing 25 μ*M* chloroquine. Transfect by adding 10 μg of DNA to 500 μl of HEPES-buffered saline (pH 7.05),[11] and then add (while vortexing) 50 μl 1.25 *M* $CaCl_2$. Immediately add this solution to the cells. At 10 hr, replace the media with 4 ml DMEM/10% FBS. It is important to remove the chloroquine in this timely fashion to prevent cell killing. We change the volume of the media to 2.5 ml (fresh media) at 36 hours post-transfection and then collect the media 24 hr later. The virus containing media should then be filtered (45 μm filter) to remove cells. The virus can now be frozen at -80°, without any significant loss of titer, or used immediately for infections.

Screening retroviral-based expression libraries for oncogenes:

Infection of recipient cell lines: We generally use the focus formation assay as the basis of our screens for transforming cDNAs.[5] These screens are performed in adherent, murine cell lines such as NIH 3T3 (fibroblasts) or C127 (epithelial) which are readily infectable by ecotropic virus, and which are susceptible to single hit transformation. If screens are to be conducted using human cell lines as recipients it would be necessary to use amphotropic packaging cell lines to generate infectious virus.

Plate 2 X 10^5 cells/plate (100 mm), 24 hr prior to infection. Immediately prior to infection, dilute the viral soup 1:1 in DMEM that contains 10% calf serum (CS) and 16 μg/ml polybrene (final concentration is 8 μg/ml). Aspirate the media from the cells and replace with the media containing virus and polybrene. Allow the infection to proceed for 3-5 hr and then replace the media with fresh DMEM/10% CS. Change the media every 2-3 days for up to 21 days and identify foci that are formed. We do not perform these screens under selective conditions since the infection frequency is usually very high, and

selection often disrupts the integrity of the monolayer, thus generating spontaneous foci. Individual foci are then scraped from the surface of the plate, transferred to individual plates (35 mm), and amplified clonally. Clonal amplification is done in the presence of hygromycin (200 μg/ml) to ensure the presence of a proviral insert. In order to determine viral titer, we remove a small aliquot of the viral soup, perform serial dilutions, and then infect recipient cells, under selection, to determine the number of resistant colonies that can be generated.

PCR amplification of proviral cDNA inserts: A PCR-mediated DNA amplification protocol has been developed to facilitate the recovery of proviral cDNA inserts (Fig. 2).[2] High molecular weight DNA, prepared from the transformed cells, is PCR amplified using a set of vector primers that flank the cDNA insert and the adjacent *supF* gene. *supF* encodes suppressor tRNA activity and can be used as a selectable marker in bacterial strains that contain the p3 plasmid. PCR products that contain the *supF* sequences are then cloned into a retroviral vector that lacks a bacterial selectable marker and selected based on their acquired suppressor tRNA activity. The PCR process utilizes *Pfu* DNA polymerase (Stratagene) which, unlike more common versions of *Taq*, has a proofreading function associated with it. Thus, errors introduced into the DNA sequences by the polymerase are kept to a minimum.

Add 600 μl lysis buffer (1 m*M* EDTA, 1% SDS, 100 μg/ml Proteinase K added directly before use) to a 35 mm well of confluent cells and incubate at 37° for 90 min. Transfer the lysate to a 1.5 ml tube and extract twice with an equal volume of phenol:chloroform. Precipitate the DNA twice with 1/10 volume 10 *M* NH_4OAc and 2

volumes 95% EtOH. Wash thoroughly with 70% EtOH after each precipitation. Dry the pellet and resuspend in 150 μl dH_2O. Prepare PCR reactions containing: 100 ng genomic DNA template, 100 ng each primer, 2.5 μl dNTP mix at 2 m*M* each, 2.5 μl *Pfu* polymerase 10X buffer, 2.5 U *Pfu* DNA polymerase, and water to 25 μl. We use the following primers for amplifying pCTV3 derived proviruses: pCTV-5' CCT CAC TCC TTC TCT AGC TC, pCTV-3' TCG AAT CAA GCT TAT CGA TAC G. PCR cycles are 95°, 60 sec; 50°, 30 sec; 68°, 6 min; X 30 cycles. At this point we run 5 μl of the PCR reaction mix on a 1.2 % gel to identify any bands that have been amplified (often there is more than one). We then do a Southern blot with a probe specific for the *supF* gene. This allows us to distinguish bands that are derived from *bona fide* proviral inserts, and allows us to identify amplified products that are present in subvisual amounts.

Cloning PCR products: pCTV3K is a specialized retroviral vector that has been developed for use in cloning PCR amplified proviral inserts (Fig. 2).[2] The vector is derived from pCTV3B and contains a Kan^r marker in place of *supF*. Ligations are performed by replacing the Kan^r marker of pCTV3K with the amplified transforming cDNA and its linked *supF* gene. Since pCTV3K does not contain any ampicillin or tetracycline resistance, recombinants can be recovered in p3 hosts with no background derived from religated vector.

Bring the volume of the PCR reaction to 180 μl with water and then add 20 μl 10 *M* NH_4OAc. Extract once with 200 μl phenol/chloroform and then precipitate with two volumes 95% EtOH. Centrifuge for 30 min at low temperature (15°) and rinse the pellet with 70% EtOH. Dissolve the pellet in water, and digest with 10 U each of *Mlu*I (37°)

and *Bsi*WI (55°). Run the digests on a 1.2% agarose gel and purify the proviral fragments (determined by the Southern blot) by electroelution and ethanol precipitation. Combine 20 ng of the PCR fragment with 40 ng of pCTV3K that has been digested with *Mlu*I and *Bsi*WI, and ligate for 1 hr at room temperature. Use the ligation mix to transform chemically competent (>5X10^6) MC1061/p3. Plasmids can be prepared by alkaline lysis procedure,[12] followed by RNase digestion. The plasmid DNA that is recovered is a fully reconstituted retroviral vector that can be used for functional testing of the transforming cDNA.

Concluding remarks:

It is our experience that these screens most often fail because of a lack of diligence in the cDNA recovery process. Because of the high efficiencies associated with retroviral infections, it is not uncommon to get multiple proviral inserts in a single transformed cell clone. Although the smaller, more abundant, PCR products are easiest to recover, it is often the larger, subvisual bands that are transforming. Such bands, if identified by Southern blot, can be readily cloned into pCTV3K. Occasionally, it is not possible to recover a transforming cDNA from a transformed clone. This will occur if transformation is due to a spontaneous event, insertional mutagenesis, or the combined expression of two or more cDNAs. Since the system has not been designed to characterize such transforming events, we have generally found it prudent to set these clones aside. If the screens are performed correctly, and carefully, a large proportion of transformed cell clones will yield transforming cDNAs.

Figure Legends:

Fig. 1. Structure of the pCTV3B retroviral-based expression cloning vector. pCTV3B consists of: (i) a 5' MMLV LTR with an extended *gag* region (Ψ) that lacks the normal initiation codon, (ii) a cDNA cloning site consisting of two *Bst*X1 sites separated by a 400 bp stuffer fragment, (iii) a *supF* gene that provides a bacterial selectable marker, (iv) the hygromycin phosphohydrolase gene that provides a mammalian selectable marker (Hygro^r), (v) a composite MPSV/MMLV 3' LTR, (vi) the replication origin from pUC8 to facilitate propagation in bacterial cells, and (vii) the simian virus 40 (SV40) origin of replication. Following viral transmission, the region between (and including) the 5' and 3' LTRs becomes stably integrated into the host genome as a proviral insert. Both the 5' and 3' LTRs of the provirus are derived from the composite MPSV-MMLV 3' LTR.

Fig. 2. Procedure for recovering transforming cDNAs from proviral inserts. Genomic DNA is isolated from transformed cell clones, and then used as a template for PCR amplification using a set of vector primers that flank the cDNA insert and the linked *supF* gene. The amplified fragment is cut with *Mlu*I and *BsiW*1 and cloned into complementary sites within the pCTV3K vector. Recombinants are isolated based on the acquisition of the suppressor tRNA activity (*supF*).

References:

[1] T. Kitamura, M. Onishi, S. Kinoshita, A. Shibuya, A. Miyajima, and G. P. Nolan, *Proc. Natl. Acad. Sci. (U S A)* **92,** 9146 (1995).

[2] I. Whitehead, H. Kirk, and R. Kay, *Mol. Cell. Biol.* **15,** 704 (1995).

[3] T. Tsukamoto, T. Huang, R. C. Guzman, X. Chen, R. V. Pascual, T. Kitamura, and S. Nandi, *Biochem. Biophys. Res. Commun.* **265,** 7 (1999).

[4] G. J. Hannon, P. Sun, A. Carnero, L. Y. Xie, R. Maestro, D. S. Conklin, and D. Beach, *Science* **283,** 1129 (1999).

[5] G. J. Clark, A. D. Cox, S. M. Graham, and C. J. Der, *Methods Enzymol.* **255,** 395 (1995).

[6] B. Seed, *Nucleic Acids Res.* **11,** 2427 (1983).

[7] H. U. Bernard, G. Krammer, and W. G. Rowekamp, *Exp. Cell. Res.* **158,** 237 (1985).

[8] M. A. Bender, T. D. Palmer, R. E. Gelinas, and A. D. Miller, *J. Virol.* **61,** 1639 (1987).

[9] B. Seed, *Nature* **329,** 840 (1987).

[10] B. Seed, and A. Aruffo, *Proc. Natl. Acad. Sci. (U S A)* **84,** 3365 (1987).

[11] T. Maniatis, E. F. Fritsch, and J. Sambrook, "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Lab., Cold Spring Harbor, NY, 1989.

[12] H. C. Birnboim and J. Doly, *Nucleic Acids Res.* **7,** 1513 (1979).

[13] W. S. Pear, G. P. Nolan, M. L. Scott, and D. Baltimore *Proc. Natl. Acad. Sci. (U S A)* **90,** 8392 (1993).